REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)
10 Nov 2001

2. REPORT TYPE
Final

3. DATES COVERED (From – To)
NA

4. TITLE AND SUBTITLE

Understanding Joint Warfighting Experiments

5a. CONTRACT NUMBER
NA

5b. GRANT NUMBER
NA

5c. PROGRAM ELEMENT NUMBER
NA

6. AUTHOR(S)

Richard A. Kass

5d. PROJECT NUMBER
NA

5e. TASK NUMBER
NA

5f. WORK UNIT NUMBER
NA

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) United States Joint Forces Command (JFCOM), Joint Experimentation Directorate (J9)
1562 Mitscher Ave, Suite 200; Norfolk VA. 23551-2488

8. PERFORMING ORGANIZATOIN REPORT NUMBER
NA

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) United States Joint Forces Command (JFCOM), Joint Experimentation Directorate (J9)
1562 Mitscher Ave, Suite 200; Norfolk VA. 23551-2488

10. SPONSOR/MONITOR'S ACRONYM(S)
USJFCOM

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
NA

12. DISTRIBUTION/AVAILABILITY STATEMENT

Available to public.

13. SUPPLEMENTARY NOTES

NA

14. ABSTRACT In October 1998 the United States Joint Forces Command (USJFCOM) established a Joint Warfighting Experimentation program to support Joint Concept Development. Experimentation is the unique scientific method for establishing whether hypothesized concepts are causally related to warfighting effects. The strengths and weaknesses of different warfighting experiment venues –constructive simulation experiments, human-in-the-loop simulation experiments, war game experiments, and field experiments—can be described in terms of their ability to meet three requirements for a successful experiment: detecting a change in the effect, identifying the cause of the change, and the ability to relate the change to real operations. A concept experimentation strategy can capitalize on the inherent strengths of these different experiment venues during different stages of concept development.

15. SUBJECT TERMS warfighting experiments, military experiments, experiment methodology, experiment validity, concept development and experimentation, joint experimentation.

16. SECURITY CLASSIFCATION: UNCLASSIFIED

b. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
UNCLASS	UNCLASS	UNCLASS	UL	11	Richard A. Kass

19b. TELEPHONE NUMBER (include area code)
(757) 836-2819

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. 239-18
298-102

20020904 064

Understanding Joint Warfighting Experiments

Richard A. Kass
United States Joint Forces Command (USJFCOM)
1562 Mitscher Ave, Suite 200
Norfolk, VA 23551-2488

(757) 836-2819
fax (757) 836-2885
kassr@je.jfcom.mil

69th MORS Symposium
Working Group P1: Warfighting Experimentation
12-14 June 2001
(November 10, 2001)

Abstract

In October 1998 the United States Joint Forces Command (USJFCOM) established a Joint Warfighting Experimentation program to support Joint Concept Development. Experimentation is the unique scientific method for establishing whether hypothesized concepts are causally related to warfighting effects. The strengths and weaknesses of different warfighting experiment venues –constructive simulation experiments, human-in-the-loop simulation experiments, war game experiments, and field experiments—can be described in terms of their ability to meet three requirements for a successful experiment: detecting a change in the effect, identifying the cause of the change, and the ability to relate the change to real operations. A concept experimentation strategy can capitalize on the inherent strengths of these different experiment venues during different stages of concept development.

Introduction

On 15 May 1998 William A. Cohen, Secretary of Defense, designated the United States Commander-in-Chief of the Atlantic Command (USCINCACOM) as the executive agent for joint warfighting experimentation within the Department of Defense.

"This effort will enable USCINCACOM to explore new joint warfighting concepts and capabilities and determine the doctrine, organization, training and education, materiel, leadership, and personnel (DOTMLP) implications for change. These experiments will support the Chairman of the Joint Chiefs of Staffs Joint Vision 2010 (JV2010) and future CJCS joint warfighting visions.[1]"

The Commander in Chief of the Atlantic Command created a new Joint Experimentation Directorate (J9) on 1

October 1998, led by a 2-star general or flag officer. The Joint Experimentation Directorate develops new joint warfighting concepts through joint experimentation for the Department of Defense. The following year the Atlantic Command was redesignated the United States Joint Forces Command (USJFCOM). This paper provides an introduction to the experimentation process and how it supports the development of joint warfighting concepts.

Experiments and Science

In 400BC Socrates, Plato, and Aristotle were among the first to investigate the meaning of knowledge and how to obtain it. Their method was primarily a rational, deductive process. Early empirical inductive methods by scientists, such as Ptolemy and Copernicus, focused on precise observations and explanations of the stars. They were not experimenters. When scientists turned from the heavens to investigate earthly objects, new and exciting answers to questions about objects within their reach could be obtained because earthly objects could be manipulated. In the early 1600s Francis Bacon introduced the term experiment and Galileo Galilei conducted experiments by rolling balls down an inclined plane to describe bodies in motion. The realization that manipulating objects would yield new knowledge spawned a new research paradigm that could not have been considered in the previous 2000 years of exploring the out-of-reach heavens. The new paradigm, called experimentation, was based on a new research question: "If I do this, what will happen?"[2] The key to understanding experimentation, and the characteristic that separates experimentation from all other research methods, is manipulating something to see what happens (Figure 1). The scientific aspect of experimentation is the manipulation of

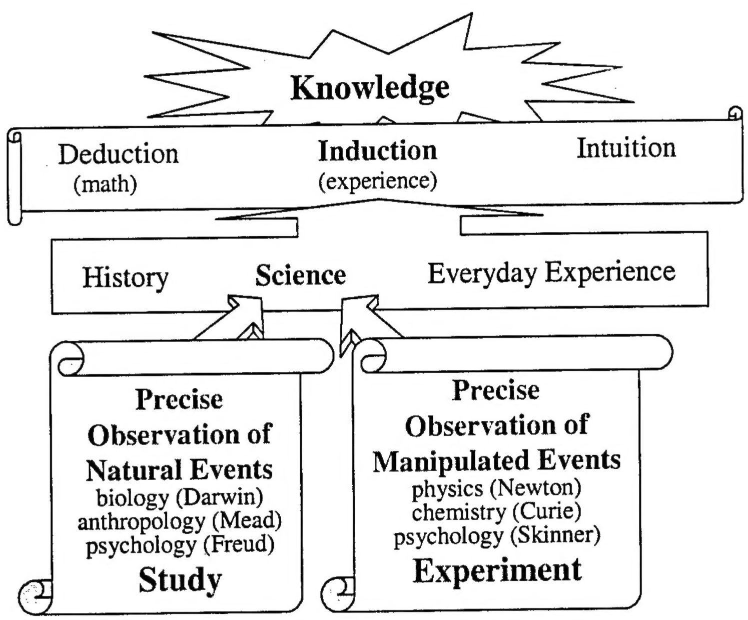

Figure 1. Experiments and knowledge.

objects under controlled conditions while taking precise measurements.

• ***Experiment*** *- test of a hypothesis under controlled conditions to substantiate and quantify the cause of an effect.*

Experiments and the Scientific Method

The scientific method has evolved during the last 400 years. The joint wafighting experiment process can be envisioned as progressing through eight steps of the scientific method (Figure 2). The process begins with the identification of a joint warfighting problem. The problem may be derived from the Joint Staff, Defense Planning Guidance, or other sources. The joint concept development process researches the joint problem and proposes (hypothesizes) possible approaches for solutions. These potential concept approaches are investigated in a warfighting experiment and the results are fed back to the concept developers to ascertain the impact of experiment results. If the results [illegible]e, such that one cannot determine if the original concept was either supported or not, then a better experiment should be designed. Clear results, on the other hand, whether positive or negative, provide an empirical basis for concept developers to improve their concept.

Experiments and Joint Warfighting Concept Development

Experimentation is not just about providing data to concept developers. It is about providing a unique type of data--data about cause and effect. New warfighting concepts are potential causes of warfighting effectiveness.

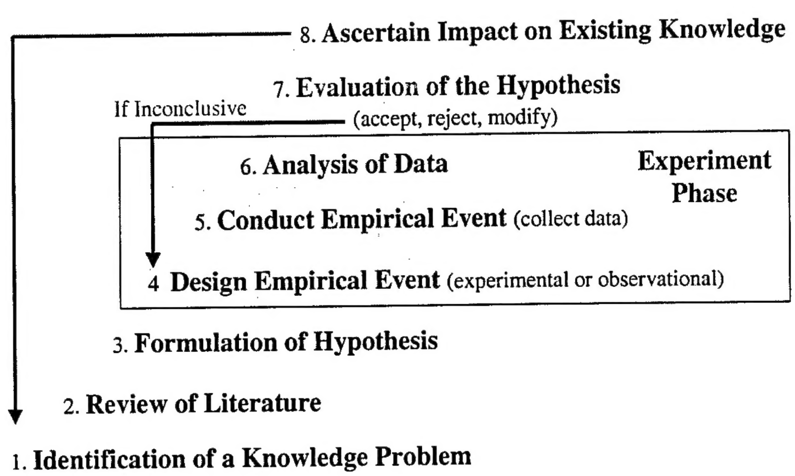

Figure 2. Scientific method.

• ***Joint Warfighting Experiment*** *– application of scientific experimentation procedures to assess the effectiveness of proposed (hypothesized) joint warfighting concept elements.*

• ***Purpose of a Joint Warfighting Experiment****—ascertain whether elements of a joint warfighting concept cause changes in military effectiveness.*

Studies and analyses and the lessons-learned process are only able to speculate on the cause of military effectiveness. Data garnered from fact-finding studies may precisely describe the tactics, organizational, personnel, and materiel resources employed in a military campaign (potential causes). Studies may also provide precise quantification of damage assessments (known effects). However, all of the studying and analysis in the world will only provide speculations (hypotheses) on which of the potential causes were truly instrumental in producing the recorded effects. Warfighting concepts are potential "causes" of military effectiveness "effects" and experimentation is the only method to empirically resolve cause and effect questions.

• ***Joint Warfighting Doctrine*** *– describes how the joint force fights today. "Fundamental principles that guide the employment of forces of two or more Services in coordinated actions towards a common objective." (JP 1-02)*

• ***Joint Warfighting Concept*** *– describes how the joint force might fight in the future.*

• ***Joint Warfighting Concept Development*** *– a progression from an initial idea for future joint warfighting to more coherent, empirically validated principles that are implemented through Doctrine, Organization, Training, Materiel, Leadership, Personnel, and Facility (DOTMLPF) recommendations.*

Experiments are required throughout the entire concept development process (Figure 3).[3] Experiments provide an empirical method for exploring, refining, and validating new ideas during the concept

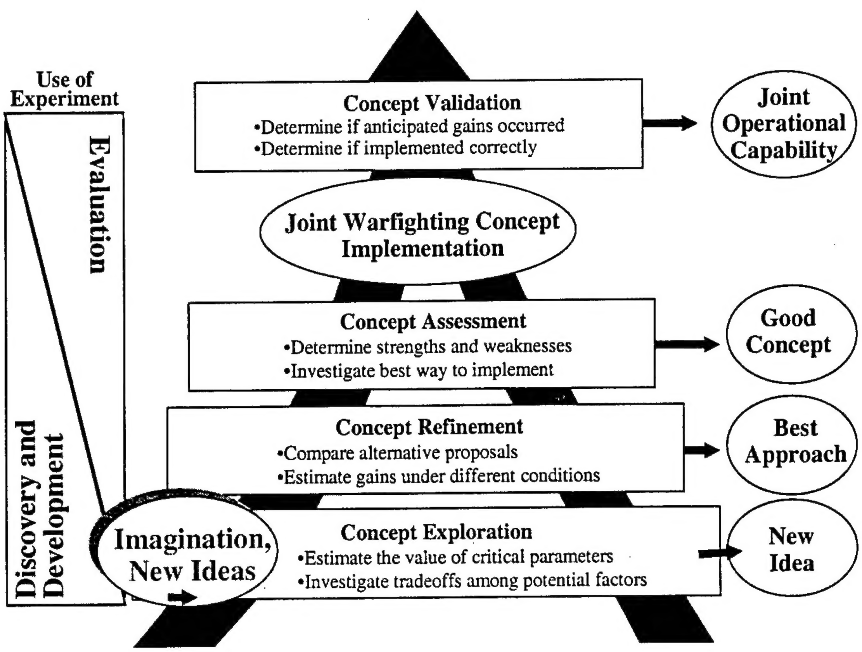

Figure 3. Designing experiments to support concept development.

development process. Early in the concept formulation cycle, experiments are used to discover and explore new ideas by determining what will be important to include in a new concept. As the concept is refined, it becomes important to investigate all proposed solutions over the full spectrum of military operations prior to focusing on a single solution. When a concept is ready for implementation, experiments can assess the proposed implementation version for strengths and weaknesses in the operational environment. A final field experiment may be used to validate the predicted gains in effectiveness in the operational force.

Language of an Experiment

Identifying experiments with the investigation of causality between warfighting concepts and warfighting effectiveness is a useful construct for organizing the language to describe the experimentation process. Any joint concept problem statement can be translated into a cause and effect question expressed as "*does A cause B?*" An experimental joint concept, a new way of doing business, can be developed for experimentation to determine if the experimental concept (A) causes a greater military effect (B). The *experiment hypothesis* provides an expectation concerning the causal observation to be observed in the experiment and is written as an *If...then...* statement with the proposed cause (concept element) identified with the *if* segment and the possible outcome (problem resolution) identified with the *then* segment.

Experiment Hypothesis
If... **"proposed DOTMLPF change"**
Then... **"improved warfighting capability"**

Components of an Experiment

All experiments—large or small, field or laboratory, military or academic, applied or pure--consist of five components (Figure 4):[4]

(1) The ***treatment***, the possible cause (A), is an element of a joint concept that is expected to influence warfighting effectiveness.

(2) The possible ***effect*** (B) of the treatment is the results of the trial, an increase or decrease in some aspect of warfighting effectiveness.

Experiment Component
Treatment: **proposed DOTMLPF change**
Effect: **improved warfighting capability**

(3) The ***experimental unit*** executes the possible cause and produces an effect.

(4) The ***trial*** is one observation of the experimental unit under Treatment A to see if effect B occurred or not and includes all of the contextual conditions under which the experiment is executed.

(5) The ***analysis*** phase of the experiment compares the results from one trial to a different trial.

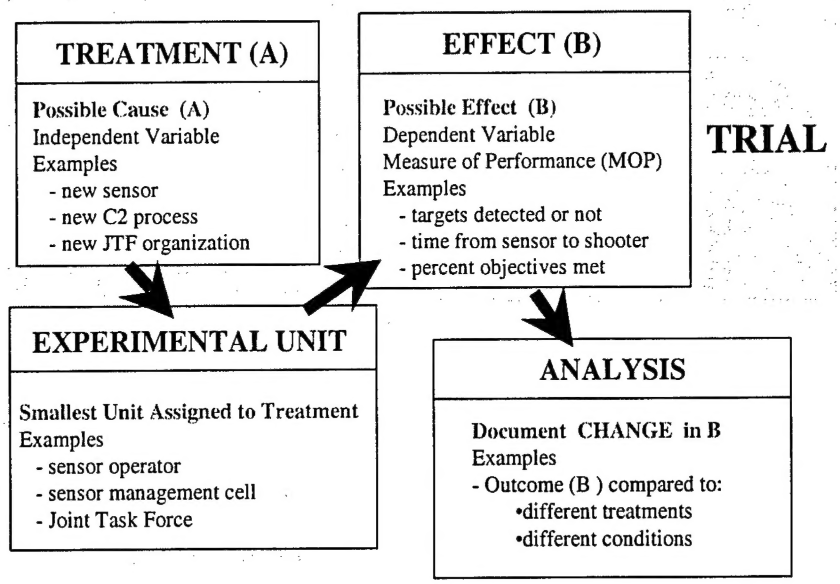

Figure 4. Five components in every experiment.

Some service field experiments are grand exercises with multiple experimental initiatives (possible causes), sometimes as many as 30 to 50 different initiatives in one experiment. The five components are useful in understanding these large field experiments. These field exercises are multiple small experiments inside the overarching experiment. Each individual experimental initiative is configurable as a unique subset of the five components. Each initiative is a separate treatment with its own experimental unit (operators in one area of a command post), its own set of outcome measures, and its own set of trial conditions which may or may not impact the other initiatives in the grand experiment. Moreover, each initiative with its five components will probably have a different number of trials.

What is a Good Experiment?

A good experiment is an experiment that provides information for ascertaining whether A caused B.[5] The scientific term for a good experiment is "valid" experiment. Three logically sequenced requirements must be met to achieve a valid experiment (Figure 5). A simple example will illustrate these three requirements. Suppose a proposed joint concept postulates that new sensors will be required to detect time critical targets. One experiment to examine this proposition might be a two-day military exercise where the old array of sensors is employed on the first day and a new sensor suite is used on day two. The primary measure of effectiveness is the percent of targets detected. The hypothesis is "If new sensors are employed, then time-critical target detections will increase." This experiment is designed to determine if the new sensors (A) will cause an increase in detections (B).

Requirement	Evidence for Validity	Threat to Validity
1 ability to **detect change**	B changed as A changed	Too much noise, can not detect any change
2 ability to **identify cause** of change	A alone caused B	Alternate explanations of change available
3 ability to **relate results** to actual operations	Change in B is expected in actual operations	Observed change may not be applicable

Valid Experiment: Evidence that A causes B

Figure 5. Three requirements for a good experiment.

Ability to detect a change in the effect (B). In the ideal situation, transition from the old to the new sensors is accompanied by a change in the percent of detections observed. If this does not occur, the concern is too much experimental noise. The ability to detect change is a signal-to-noise problem. Too much experimental error produces too much variability, making it difficult to detect a change. Many experiment techniques are designed to reduce experiment variation: calibrating instrumentation to reduce data collection variation, controlling stimuli (the targets) presentations to only one or two variations to reduce response (detections) variation, and controlling the external environment (time of day, visibility, etc). Sample size is another consideration for reducing the signal-to-noise ratio. The computation of variability in statistics decreases with the square of the number of observations.

To detect *change,* experiments require two or more trials: before and after treatment, various treatment levels, alternate competing treatments, or the same treatment under different conditions. When large-scale military field exercises are used to support an experiment, resources often inhibit multiple field trials. The value of a single field exercise is enhanced within a model-exercise-model experiment paradigm. Multiple trials with multiple iterations are executed in the pre-exercise simulation. The field exercise is designed to replicate the key conditions and results from the pre-simulation results. The data from the exercise are then used to calibrate the post-exercise re-simulation of the experiment trials, yielding greater confidence in the simulated results.

Ability to identify the cause (A) of change. Lets suppose the experimenter had a good experimental design that reduced

variability and produced a change (increase) in the percent of detections. The question now is whether the detected change was due to the intended cause, changing from old sensors to new, or due to something else. The scientific term for alternate explanations of experimental data is *confounded* results. In this example an alternate explanation for the increase in detections on day two is that it was due to a learning effect. The sensor operators may have been more adept at finding targets as a result of their experience with target presentations on day one and, consequently, would have increased target detections on day two whether the sensors were changed or not. This would dramatically change the conclusion of the detected change.

Scientists have developed experimental techniques to eliminate alternative explanations of the cause of change. These include counter balancing the presentation of stimuli to the experimental unit, the use of placebos in drug research, use of a control group, randomizing participants between treatment groups, and elimination or control of external influencers.

Ability to relate the results to actual operations. Again, let's suppose that the experimenter was successful in detecting change and isolating the cause. Now the question is whether the experimental results are applicable to the operational forces in actual military operations. Experimental results are only useful to the extent they say something about the real world. *Generalizability* is the scientific term for the ability to apply results outside the experiment context. Ability to relate results pertains to experiment realism. Experimental design issues supporting operational realism revolve around the representation of surrogate systems, the use of operational forces as the experimental unit, and the use of operational scenarios with a realistic reactive threat.

Different Warfighting Experimentation Methods Provide Different Strengths

All experiments are a balance between the three requirements discussed above. Attempts to satisfy one of the requirements works against satisfying the other two. Consequently, 100-percent valid experiments are not achievable. Precision and control increase the ability to detect change and isolate the cause but decrease the ability to apply the results to real-world situations because real-world operations are not precise. Experiments designed to detect and identify change emphasize strict control of trial conditions and multiple repetitions of similar events. On the other hand, experiments designed to relate results emphasize free-play, uncertainty, and a reactive threat.

Since one experiment cannot satisfy all three experiment requirements equally, different requirements are emphasized in any given experiment (Figure 6). Experiments are designed to consider requirement tradeoffs and are designed to minimize the loss of one requirement because of the priority of another. Requirement tradeoffs are inherent in the four different warfighting experiment methods.

Wargame experiments typically employ command and staff officers planning a military campaign. At certain decision points the blue players' plans are given to a neutral, white cell, who then allows the red players to plan a counter move, and so on. Each move is arbitrated by the white cell and they might use a simulation to assist in

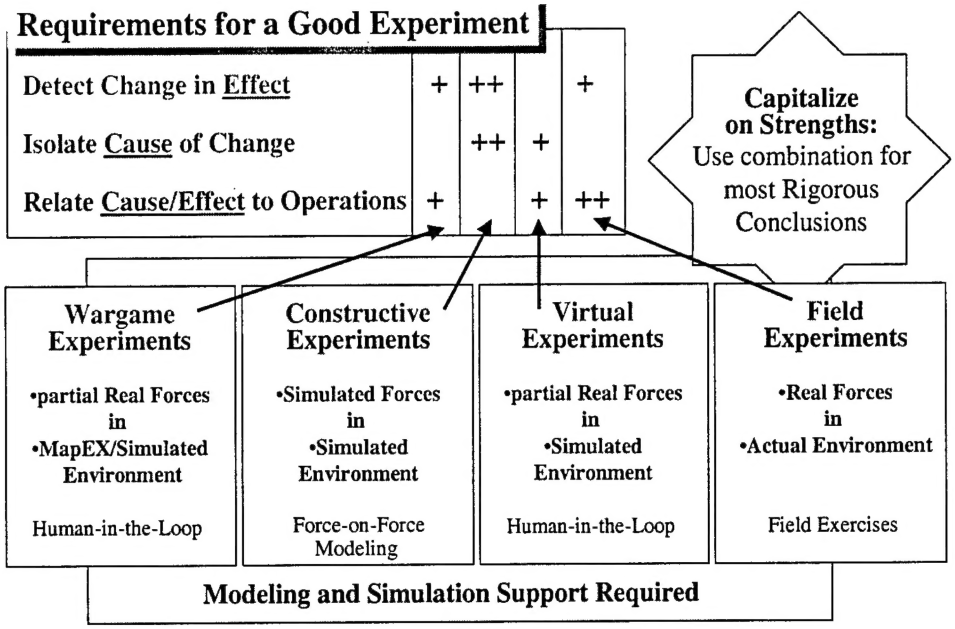

Figure 6. Relative strengths of different warfighting experimentation methods.

determining the outcome of each move. An experiment built on the wargame methodology might involve fighting the same campaign using two or more different strategies. The strength of wargame experiments resides in the ability to detect any change in the wargame outcome provided there are major differences in the strategies used. Additionally, to the extent operational scenarios and existing commands are used as players, wargame experiments may reflect real-world possibilities. A major limitation is the inability to isolate the true cause of change because of the myriad of differences between playing two different campaigns against a reactive threat. ***Constructive experiments*** reflect the closed-loop force-on-force simulation employed by the model and simulation community. "Closed-loop" indicates no human intervention in the play of the simulation once the initial parameters are chosen and the simulation is started and run to completion. Constructive simulations are the mainstay of warfighting experimentation and are employed by all military analytical agencies. Constructive simulations allow repeated replay of the same battle under identical conditions while systematically varying parameters between trial runs: a new weapon or sensor characteristic, different resource, different tactic, or different threat. Constructive experiments with multiple runs are ideal for detecting change and isolating the cause of that change. Unfortunately, constructive simulations are often questioned with respect to the applicability of results to operational situations due to the number of assumptions that must be made in the attempt to model complex events.

Virtual experiments are between pure constructive experiments and field

experiments. Virtual experiments employ human-in-the-loop simulations. The prototype virtual simulation is the flight simulator. The human pilot makes all the decisions and controls the real-time inputs while the simulation provides artificial, yet realistic real-time feedback. In a command and control virtual simulation experiment, a sensor operator might receive real-time simulated sensor inputs and makes real-time decisions to launch simulated weapons against simulated targets. The use of actual military operators allows this type of experiment to better reflect warfighting decision making than the pure closed-loop constructive experiments. However, once human decisions are introduced, variability increases making it more difficult to detect changes.

Field experiments are wargames conducted in the actual environment with actual military units and equipment. As such, these experiments have the highest applicability of results to real situations. Good field experiments, like good military exercises, are the closest thing to real military operations. Because field experiments include much of the uncertainty, variability, and challenges of actual operations, the ability to isolate the true cause of any detected change will suffer.

Emphasizing Different Experiment Requirements During Concept Development

Since no single experiment can meet all three experiment requirements to the fullest; and since different experiment methods emphasize the three requirements differently, a comprehensive experimentation program needs to capitalize on the strengths of each method (Figure 7).

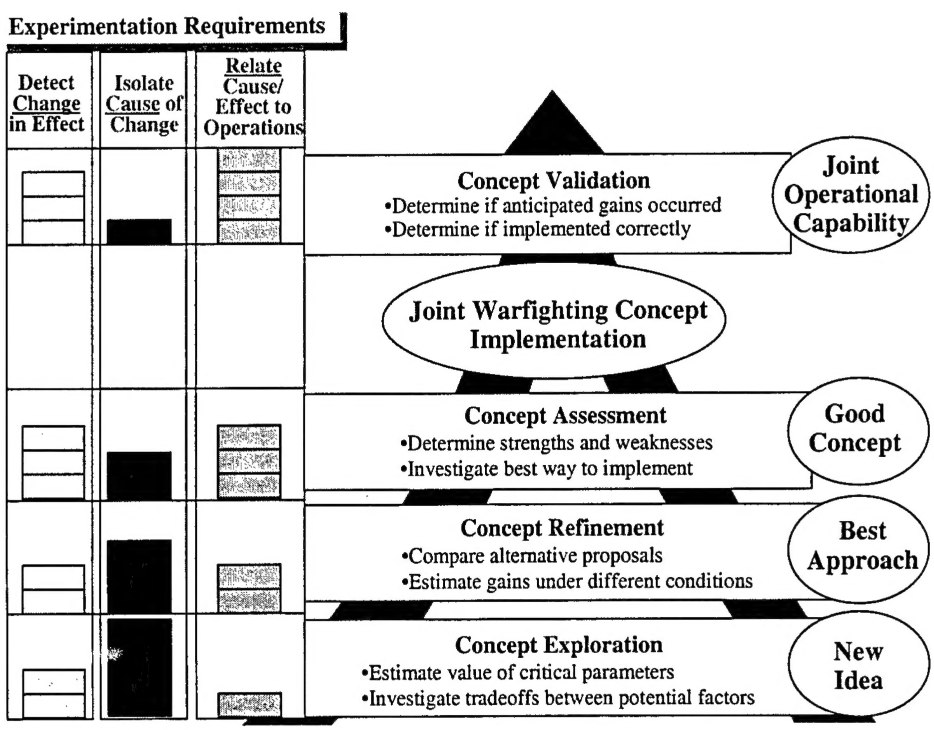

Figure 7. Relative importance of the three experiment requirements during the concept development process.

Where one expects a small effect and it is important to determine the precise relationship between the treatment and its effect, the priority should be detecting change and correctly isolating the cause of that change. On the other hand, if one expects a large effect, and if it is important to determine whether the effect will occur in the operational environment with typical units, and there is less need to address questions of why the specific result occurred, then ability to apply results is the priority. These guidelines can be considered for selecting experimentation methods at different stages in the concept development cycle.

Early in the concept development cycle, when attempting to synthesize and understand a new warfighting concept, it is important to examine the relationships among the components to determine which components are essential. It is important to accurately identify possible component causes at this early stage to avoid spending resources on those elements that will not pan out as causes and, conversely, avoid eliminating those with future potential.

In later concept development phases, when alternative concepts are examined under different conditions, it becomes important to focus more attention to the applicability of the results to different scenarios and conditions. After the concept has been narrowed to the best alternative, more emphasis in the experiment is placed on realistic conditions and scenarios and whether or not a large enough change is detected to warrant implementation. During the final field evaluation phase the emphasis is on validating the expected gain in warfighting effectiveness in operational forces.

Summary

Joint warfighting experiments are essential to developing empirical-based concepts. Joint concepts describe the doctrine, organization, training, materiel, leadership, personnel, and facilities that will enable or cause future joint warfighting effectiveness. Experimentation is the unique scientific method for establishing whether hypothesized concepts are causally related to effects. If the five experiment components are designed to meet the three experiment requirements, the warfighting experiment will provide the concept developer with the basis to proceed. This "develop-experiment-refine" concept development process ensures that new joint warfighting concept will be related to warfighting effectiveness; thus providing the foundation for transforming the United States military forces.

---

[1] William A. Cohen, Secretary of Defense, Charter for Joint Experimentation, 15 may 1998.

[2] From Richard P. Feynman's description of science in The Meaning of It All: Thoughts of a Citizen Scientist, Helix Books, 1998.

[3]Adapted from Riecken, H.W. & Boruch, R.F. Social Experimentation: A Method for Planning and Evaluating Social Interventions, Academic Press, 1974.

[4] From Cook, D.T.& Campbell, D.T. Quasi-Experimentation: Design and Analysis Issues for Field Settings, Rand McNally, 1979. For application of these concepts to test and evaluation, see Kass, R.A. "Design of Valid Operational Tests," International Journal of Test and Evaluation, June/July, pages 51-59, 1997.

[5] Requirements based on Cook and Campbell (1979) and the earlier work by Campbell, D.T. & Stanley, J.C. Experimental and Quasi-experimental Designs for Research, Rand McNally, eighth printing, 1972.